

**Н. В. ВЛАСЕНКО**, асп. ХНУРЭ, Харьков;  
**А. С. КАНУННИКОВ**, студент ХНУРЭ, Харьков;  
**С. В. МАШТАЛИР**, канд. техн. наук, доц. ХНУРЭ, Харьков

## КОМПРЕССИРОВАНИЕ ОПИСАНИЯ ВИЗУАЛЬНЫХ ОБЪЕКТОВ ПУТЕМ ФИЛЬТРАЦИИ ЕГО КОМПОНЕНТ ПО КРИТЕРИЮ СТАБИЛЬНОСТИ

Досліджуються шляхи стиснення простору структурних ознак зображень із використанням критерію стабільності їх значень на навчальній множині геометричних перетворень. Стиснений опис формується на основі обчисленої підмножини стабільних точок. Результати експериментів по класифікації зображень на основі запропонованих методів підтверджують їх високу ефективність.

**Ключові слова:** класифікація зображень, характерні ознаки, структурний опис, стиснення, інваріантність, обчислювальні затрати.

Исследуются пути сжатия пространства структурных признаков изображений с использованием критерия стабильности их значений на обучающем множестве геометрических преобразований. Сжатое описание формируется на основе вычисленного подмножества стабильных точек. Результаты экспериментов по классификации изображений на основе предложенных методов подтверждают их высокую эффективность.

**Ключевые слова:** классификация изображений, характерные признаки, структурное описание, сжатие, инвариантность, вычислительные затраты.

Ways of compression space structural features images using the criterion of stability of their values on the training set of geometric transformations explores in article. A concise description is based on the computed subset of stable points. Experimental results on image classification based on the proposed methods confirm their high efficiency.

**Keywords:** image classification, features, structural description, compression, invariance, the computational cost.

**Введение.** Формирование системы признаков – ключевая проблема в задачах распознавания и анализа многомерных данных [1,2]. Ее решение часто является предметом компромисса относительно нескольких противоречивых требований. С одной стороны, часто желательно, чтобы значения признаков объекта были близки между собой внутри одного и того же описания (однородность). С другой стороны, величины признаков в целях успешного распознавания должны как можно больше различаться для описаний объектов из разных классов.

Механизм человеческого зрения в сравнении со зрительными системами искусственного интеллекта в высшей степени обладает возможностью выделять наиболее важную для распознавания видеoinформацию. Это могут быть как локальные характеристики (ключевые точки), так и глобальные данные (взаимное размещение частей объекта, контраст, относительные

размеры и др.). Например, в процессе слежения за объектом человеческий глаз опирается, как правило, на незначительное число устойчивых к изменениям характеристик объекта. Естественно выдвинуть гипотезу о том, что применение разрабатываемых интеллектуальных методов для систем обработки видеоданных должно быть связано с формированием незначительного по мощности, но в то же время эффективного для классификации подмножества устойчивых по некоторому критерию признаков.

Немаловажную роль при классификации изображений играет свойство инвариантности к геометрическим преобразованиям, которое определяется используемым способом формирования элементов описания (детекторы ключевых точек SURF, SIFT) [1]. Основным требованием к таким детекторам является обеспечение достаточного уровня стабильности значений вычисленных инвариантных признаков при изменении параметров геометрических преобразований.

При применении структурных методов распознавания изображений множество признаков обычно связано с пространством  $R^n$  векторов фиксированной размерности с действительными компонентами [1–3]. Структурное описание видео-объекта – это конечное множество  $Z \subset R^n$ .

Формально снижение размерности пространства признаков в общем случае сводится к построению отображения:

$$\Xi: X_1 \times X_2 \times \dots \times X_p \rightarrow Y_1 \times Y_2 \times \dots \times Y_q, \quad q < p, \quad (1)$$

где  $X_1, \dots, X_p$  – исходный набор признаков,  $Y_1, \dots, Y_q$  – новый набор признаков,  $\times$  – символ декартова произведения множеств,  $p, q$  – размерности исходного и сжатого пространств [4]. При этом считается, что каждый из признаков имеет свою область значений. Применительно к описаниям  $X_i \in R^n, Y_j \in R^n$  можно рассмотреть два основных способа построения  $\Xi$ : путем сокращения числа векторов в  $Z$  и/или уменьшения размерности  $n$  отдельного вектора, т.е. через переход к пространству  $R^v, v < n$ . В частности, построение  $R^v$  можно реализовать путем перекувантования или передискретизации вектора признаков по значению отдельной компоненты или по всему числу компонент.

Сокращение числа векторов в описании  $Z$  предполагает построение сжатого описания  $Z^*$  на основе отображения (1) вида  $\Xi: Z \rightarrow Z^*$ , вариантом которого есть формирование подмножества  $Z^* \subset Z$  значительно меньшей мощности путем применения некоторой процедуры фильтрации. В теории распознавания этот способ ассоциируется с формированием множества наиболее значимых признаков. Другой вариант построения  $\Xi$  состоит в

смысловом агрегировании элементов  $Z$  путем объединения их свойств в более глобальные структуры данных, формирующие  $Z^*$  [4]. В целом компрессирование направлено на уменьшение объема вычислений и снижение требований к методам распознавания с точки зрения быстродействия. Выигрыш во времени пропорционален отношению мощностей старой и новой (сжатой) систем признаков. В то же время надо иметь в виду, что уменьшение объема описания путем компрессирования может привести к снижению вероятности распознавания в условиях помех.

Возьмем за основу снижения размерности признакового пространства одно из важнейших свойств – устойчивость значений инвариантных признаков описания при действии геометрических преобразований. Понятно, что устойчивость в полной мере определяется используемым детектором признаков. Другие принципы фильтрации данных могут применять, например, отбор наиболее информативных элементов описания.

Цель исследования – построение метода компрессирования структурного описания на основе свойства стабильности его компонент в условиях влияния геометрических преобразований.

Задачи статьи состоят в развитии теории сжатия в признаковом пространстве с применением свойства стабильности, вычислении подобия для сжатых описаний, осуществлении экспериментальных исследований и сопоставлении результатов с другими методами для конкретных баз видеoinформации.

**Способ формирования сжатого описания.** Построим вначале множество изображений, которое будет основой установления и формирования устойчивого сокращенного подмножества признаков. Для этого выберем произвольный эталон из видео-базы и подвергнем его фиксированному числу  $N$  различных значений геометрических преобразований из группы  $G = \{g_i\}_{i=1}^N$ . Одним из вариантов обработки есть дискретизация всего допустимого в практических приложениях пространства преобразований. Например, для преобразования поворота последовательно вращаем эталон на дискретные углы  $g_1 = 20^\circ$ , и т.д. до  $g_{18} = 360^\circ$ , здесь  $N = 18$ . Еще одним приемлемым способом формирования  $G$  может быть процедура случайного выбора его  $N$  элементов по равномерному закону из некоторого конечного множества.

Для каждого из преобразованных изображений путем применения детектора формируем описание  $Z_1$  в виде множества векторов характерных признаков. Обозначим  $S$  – размер исходного эталонного описания  $Z$  (при угле поворота  $g_0 = 0^\circ$ ),  $s_1$  – размер описания  $Z_1$  преобразованного эталона.

Применим теперь отображение  $\Theta$  установления соответствий  $\Theta: Z \rightarrow Z_1$  элементов двух описаний, в процессе реализации  $\Theta$  для каждого

$z \in Z$  попытаемся установить элемент  $z_1 \in Z_1$  через построение соответствия  $\beta(z, z_1) \in \{0, 1\}$ , которое равно 1, если  $z_1$  сформирован, и 0 – в противном случае [5]. Формально соответствие  $\Theta$  между множествами  $Z, Z_1$  с элементами  $z \in Z, z_1 \in Z_1$  определено как подмножество декартова произведения  $\Theta \subseteq Z \times Z_1$  и может быть задано таблицей. С каждым из сформированных соответствий  $\beta(z, z_1)$  свяжем значение  $\rho(z, z_1)$  метрики, путем минимизации значений которой оно синтезируется. В качестве  $\rho(\cdot)$  может быть использована произвольная метрика в  $R^n$ . Выполним отображение  $\Theta$ , в результате для каждого  $z \in Z$  получим значение  $\beta$ . Сформируем бинарный массив  $\beta^g = \{\beta_1^g, \beta_2^g, \dots, \beta_s^g\}$ , который свяжем со значением преобразования  $g$ .

Проведем испытания для всех изображений, полученных путем применения к выбранному эталону геометрических преобразований из  $G = \{g_i\}_{i=1}^N$ . Для каждого номера  $q$  элемента в векторе  $\beta^g$  подсчитаем число сформированных соответствий

$$k_q = \sum_{i=1}^N \beta_q^i, \quad q \in [1, s], k_q \in [0, N] \quad (2)$$

в серии из  $N$  испытаний. Будем использовать критерий (2) как характеристику устойчивости  $q$ -го признака эталона. По результатам обработки на множестве  $G$  сформируем таблицу значений критерия  $k_q$  для всех  $s$  признаков. Для однозначных соответствий выполняется  $k_q \leq N$ . При этом факт возможного формирования соответствия с «чужим» признаком не контролируется. Правильность установления соответствия полностью определяется отображением  $\Theta$ . Значение критерия  $k_q$  отражает рейтинг устойчивости для  $q$ -го признака.

Далее обработаем элементы множества  $\{k_q\}_{q=1}^s$  путем ранжирования и выделим в нем подмножество  $Q^m$  из  $m$  номеров элементов с наибольшим значением рейтинга  $k_q$ . Еще одним вариантом обработки может быть формирование  $Q^m$  из номеров элементов, имеющих максимальное значение  $k_q$ . В этом случае величина  $m$  получается как результат анализа. Понятно, что выбор  $m$  для успешного решения задачи распознавания зависит от эталонов как элементов видео-базы. Описанный подход к формированию сжатого описания (схема на рис. 1) можно характеризовать как синтез подмножества наиболее устойчивых компонентов описания.

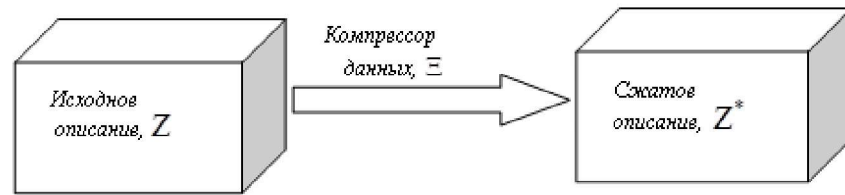


Рис. 1 – Схема компрессирования

Другим вариантом обсуждаемого способа компрессирования данных может быть использование для рейтингования компонентов критерия в виде суммы значений минимумов расстояний, при которых сформировано каждое соответствие из  $Q^m$ . Значение  $r_q$  этого критерия опишем как сумму

$$r_q = \sum_{i=1}^{k_q} \rho(z_q, z_{li}), \quad (3)$$

где  $\rho(z_q, z_{li})$  – расстояние для сформированного соответствия  $\beta_i^q = 1$ . Чаще всего в качестве  $\rho$  используют евклидову метрику. Чем меньше значение  $r_q$ , тем лучшим образом сформировано сжатое множество  $Q^m$ . В идеальном случае  $r_q = 0$ . Возможно также совместное использование критериев (2), (3) в одной процедуре. Значение критериев представим в виде таблицы.

Значения критериев устойчивости признаков

Номер признака эталона	1	2	...	s-1	s
Число соответствий	$k_1$	$k_2$	...	$k_{s-1}$	$k_s$
Сумма расстояний	$r_1$	$r_2$	...	$r_{s-1}$	$r_s$

Обработка значений таблицы в целях формирования  $Q^m$  может быть выполнена несколькими способами, такими как: а) пороговая обработка, при котором в  $Q^m$  отбираются те номера признаков, для которых  $k_q \geq \delta_k$ , где  $\delta_k$  – порог для числа повторений признака; б) ранжирование  $\{k_q\}$  по убыванию и выделение из полученного ряда  $m$  первых признаков. В любом случае для эталона должно быть сформировано не менее чем  $m$  признаков, с использованием которых будет осуществляться распознавание или классификация.

Результат формирования сжатого описания  $Q^m$  на основе сопоставления зависит от ряда факторов, в частности, от способа получения признаков (в большей степени) и от правила, устанавливающего эквивалентность

компонентов описаний. Свойства детектора и правило определяются рядом параметров. К ним относятся: порог  $\delta_p$  на величину минимума расстояния  $\rho(z_q, z_{li})$  между векторами (в традиционном варианте принято  $\delta_p = 0,24$ , что соответствует 1,5% от максимально возможного отклонения [5]), значение отношения минимума к ближайшему минимуму (принимается равным 0,7). Кроме того, на результат в значительной степени влияет также способ моделирования геометрических преобразований. Так, при применении методов интерполяции повышается число достигнутых соответствий за счет возрастания числа точек.

Рассмотренную обработку можно применить и для других преобразований, например, масштабирования, а также для преобразований комплексированного типа. В то же время для наиболее простых преобразований смещения на целые величины, где нет искажений, связанных с дискретизацией раstra, обсуждаемый подход напрямую не применим, т.к. вследствие отсутствия потерь точек описания значение  $k_q$  для всех признаков будет равно  $N$ , и стабильные точки можно выбирать произвольным образом. Для смешанных преобразований, применяемых на практике (смещения, поворот, масштаб) применение подхода не вызывает затруднений.

По результатам обобщенного анализа для всех эталонов видео-базы сформируем подмножества  $Q_i^m$  для  $i$ -го эталона. В результате достигаем сокращения объема информации в базе данных, представляя каждый из эталонов описанием из  $m$  элементов. Эта процедура допускает обобщения. Например, можно установить разное число наиболее устойчивых характерных признаков для разных эталонов по порогу для значения метрики. Другой вариант построения сжатой базы может быть связан с последовательной процедурой отбора признаков, которые не содержатся в уже синтезированных описаниях других эталонов.

**Результаты компьютерных экспериментов.** Гипотеза об объективном существовании устойчивого при геометрических преобразованиях подмножества характерных признаков (ХП) получила полное подтверждение в эксперименте. Проведение компьютерного моделирования осуществлялось для нескольких баз изображений, в частности, для изображений шахматных фигур (шесть изображений, рис. 2). Отображение (1) реализовано программно путем формирования для эталонного описания массива «стабильных» точек при осуществлении конечного числа  $N = 18$  поворотов. Множество стабильных точек формировалось как набор с наибольшим числом повторений. На рис. 3 приведены примеры эталонного изображения шахматной фигуры коня со сформированными характерными признаками (рис. 3, а, 239 точек), а также преобразованное изображение (рис. 3, б, в) с

разным числом стабильных точек при значениях параметра  $\delta_p=0,24$  (64 точки),  $\delta_p=0,16$  (31 точка). При  $\delta_p=0,08$  число стабильных точек равнялось 6. Такие значения порога отражают точность сопоставления в виде процента от максимального значения отклонения между векторами SURF в евклидовой метрике: 0,08–0,5%, 0,16–1%, 0,24–1,5%. Из рис. 3 видно, что путем изменения параметра  $\delta_p$ , фактически отражающего размер допустимой окрестности для эквивалентных векторов в  $R^n$ , можно управлять числом стабильных точек.

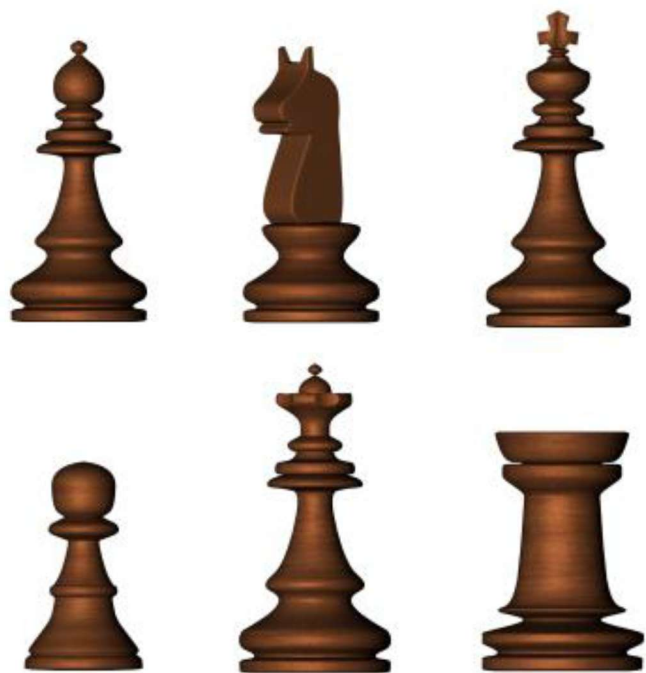


Рис. 2 – База видео-объектов

На рис. 4 приведены гистограммы для числа повторений точек (по оси абсцисс – число повторений, по оси ординат – число точек описания). Каждая гистограмма содержит 18 столбиков. Вид гистограмм рис. 4 подтверждает целесообразность проведения классификации точек по критерию устойчивости, т.к. максимумы гистограмм имеют ярко выраженный характер, особенно при малых значениях порога  $\delta_p$ . При наибольшем из исследуемых

значений  $\delta_p=0,24$  гистограмма выравнивается, и число точек с максимальным числом повторений возрастает.

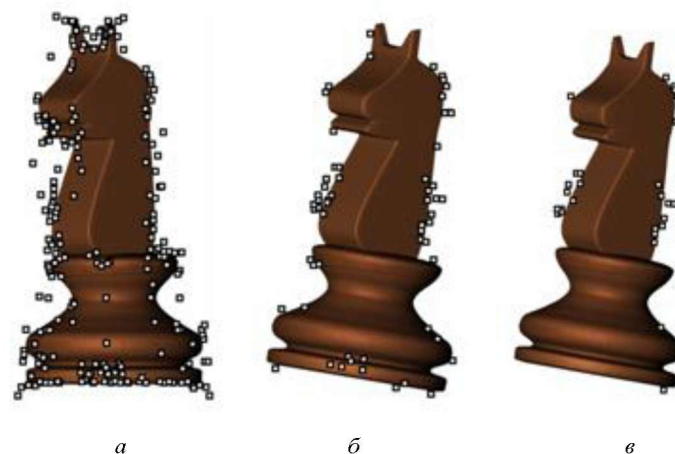


Рис. 3 – Изображения с ключевыми признаками: а – эталонное (число признаков 239), б – при  $\delta_p=0,24$  (64), в –  $\delta_p=0,16$  (31)

Анализ показывает, что достаточно большое число точек (19) имеют близкие параметры устойчивости (повторяются 16–17 раз из 18). При этом также существенное число точек (до 49) имеют очень низкий коэффициент повторения (0–3), что в целом подтверждает практическую целесообразность разделения исходного описания на подмножества, т.к. получается, что существенное число точек сформированного описания можно отнести к «мусору» при распознавании.

Таким образом, при  $m=31$  коэффициент сжатия  $\gamma$  можно оценить как  $\gamma = 239/31 \approx 7,7$ . Заметим, что полученное сжатое описание из 31 точки в соответствии с описанной схемой обработки допускает дальнейшее сокращение за счет уменьшения величины  $m$  путем исключения элементов по какому-то иному принципу. Как видим, параметром  $m$  можно управлять, непосредственно влияя на уровень компрессии описания. При  $\delta_p=0,08$  число точек с коэффициентом повторения 18 равно 6, и, соответственно, коэффициент сжатия достигает величины  $\gamma = 239/6 \approx 39,8$ .

Моделирование для преобразований масштаба при пороге  $\delta_p=0,16$  привело к формированию 9 наиболее стабильных точек описания, причем отсутствует пересечение этого множества стабильных точек с множеством точек при повороте для одного и того же изображения эталона. В то же время



при пороге  $\delta_p=0,24$  эти 9 точек уже попадают в состав множества стабильных точек. Это в очередной раз подтверждает гипотезу об объективном существовании подмножества стабильных точек описания. Результаты проведенных экспериментов для преобразований поворота и масштаба (значения масштаба 0,8, 1,0, 1,2,  $N=54$ ,  $\delta_p=0,24$ ) отображены в виде гистограммы рис. 4, *г* и демонстрируют стабильность 46 точек из 239 с максимальным числом повторений, равным  $N=54$ . Это говорит об универсальности описанного подхода для практического разнообразия допустимых преобразований.

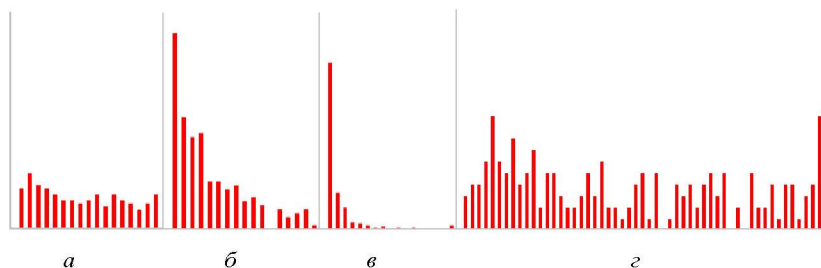


Рис. 4 – Гистограммы распределения числа повторяющихся точек при порогах  $\delta_p=0,24(a)$ ,  $\delta_p=0,16(b)$ ,  $\delta_p=0,08(v)$ ,  $\delta_p=0,24(g)$  – поворот и масштаб

Важным моментом при построении сжатых описаний для любого подхода остается сохранение высокого уровня вероятности распознавания. При этом понятно, что уменьшение размерности пространства путем фильтрации в общем случае непреодолимо приведет к снижению этой вероятности. Динамика снижения непосредственно связана с базой видеоданных, для которой осуществляется классификация. И в произвольной базе качество различения определяется самым слабым звеном – двумя наиболее схожими в пространстве объектами. Снижение вероятности не произойдет лишь в идеальной ситуации, если множества признаков для всех эталонов базы вообще не имеют пересечений. Этим свойством не обладает база шахматных фигур. Дело в том, что для отдельных фигур базы, например, короля и ферзя, число общих точек достигает значения 58% (269 общих точек из 465).

По этой причине были проведены сравнительные эксперименты для базы домашних животных [5]. В целом для неё сохраняются полученные выводы по возможности формирования подмножества стабильных точек, но число общих точек этой базы не превышает 5% от их общего количества.

Несмотря на указанное замечание, даже для базы шахматных фигур сохраняется безошибочная классификация на уровне 40% от всего числа точек описания, что допускает уровень компрессирования 2,5 раза.

**Выводы.** В результате проведенного исследования доказана гипотеза о существовании подмножества устойчивых точек в структурном описании видео-объекта в виде множества инвариантных характерных признаков. Степень устойчивости определяется применяемым методом для выделения характерных признаков, а также порогом точности при установлении их эквивалентности.

Предложенный в статье новый метод компрессирования описания видео-объекта состоит в реализации следующих шагов: 1) установление подмножества наиболее стабильных точек описания с точки зрения сохранения свойства инвариантности к группе геометрических преобразований; 2) формирование сжатого описания из состава элементов полученного подмножества стабильных компонент описания. Классификация объектов при этом выполняется в обычном режиме с использованием компрессированного описания в виде сокращенного числа точек, что значительно эффективнее с точки зрения быстродействия при решении прикладных задач. Вероятность правильной классификации при этом сохраняется на высоком уровне.

Практически важным результатом есть экспериментальное подтверждение теоретических принципов построения системы сжатия на основе формирования стабильных точек, что говорит о целесообразности применения предложенных методов в задачах компьютерного зрения. Несомненным преимуществом подхода есть существенный выигрыш во времени обработки и сокращение объема описания. Перспективы исследования состоят в обобщении подхода в плане его универсального применения для произвольной базы видеоданных с точки зрения получения предельного числа точек эталонов для осуществления качественной классификации.

**Список литературы.** 1. Шати́ро Л. Компьютерное зрение / Л. Шати́ро, Дж. Стокман, пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 2. Прикладная статистика: Классификация и снижение размерности: Справ. изд. / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин; под ред. С. А. Айвазяна. – М.: Финансы и статистика, 1989. – 607 с. 3. Гаганов В. Инвариантные алгоритмы сопоставления точечных особенностей на изображениях [Электронный ресурс] / В. Гаганов. – Режим доступа: [http://cgm.computergraphics.ru/issues/issue17/invariant\\_features](http://cgm.computergraphics.ru/issues/issue17/invariant_features). 4. Петровский А. Б. Снижение размерности признакового пространства в задачах многокритериальной классификации / А. Б. Петровский, Г. В. Ройзензон // Decision Making and Business Intelligence Strategies and Technique. – Ed. by K. Markov, K. Ivanova, I. Mitov. – No. 3. – Sofia: FOI ITHEA. – 2008. – P. 81–86. 5. Власенко Н. В. Анализ кортежей наиболее значимых соответствий между описаниями объектов в задачах метрического распознавания / Н. В. Власенко, В. П. Маисталир // Реєстрація, зберігання і обробка даних. – 2012. – Т. 14, № 3. – С. 20–32. 6. Duda R. O. Pattern classification / Duda R. O., Hart P. E., Stork D. G. – 2 ed., Wiley, 2000. – 738 p.

Надійшла до редколегії 08.12.2012